



# UNITED STATES PATENT AND TRADEMARK OFFICE

*mn*  
UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/768,178	02/02/2004	Xiaohong Huang	122.1580	3651
21171	7590	06/11/2007		
STAAS & HALSEY LLP SUITE 700 1201 NEW YORK AVENUE, N.W. WASHINGTON, DC 20005			EXAMINER DEBROW, JAMES J	
			ART UNIT 2176	PAPER NUMBER
			MAIL DATE 06/11/2007	DELIVERY MODE PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	Application No. 10/768,178	Applicant(s) HUANG ET AL.	
	Examiner James J. Debrow	Art Unit 2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) ☒ Responsive to communication(s) filed on 15 March 2007.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) ☒ Claim(s) 1-14 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-14 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
a) ☐ All    b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)                                | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                       | 5) <input type="checkbox"/> Notice of Informal Patent Application                       |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____  |

### DETAILED ACTION

1. This action is responsive to communications: Amendment filed 15 Mar. 2007.
2. Claims 1-14 are pending in this case. Claims 1 and 7 are independent claims.

### *Applicant's Response*

3. In Applicant's Response dated 15 Mar. 2007, Applicant amended claims 1-12; added new claims 13 and 14. Applicant argued against all rejections previously set forth in the Office Action dated 15 Dec. 2006.

### *Claim Rejections - 35 USC § 102*

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

5. **Claims 1, 2, 7, 8, 13 and 14 are rejected under 35 U.S.C. 102(b) as being anticipated by Maeda et al. (Pub. No.: US 2006/0004780 A1; effective Filing Date: Jun. 29, 1999) (hereinafter "Maeda").**

**In regards to independent claim 1, Maeda discloses *an apparatus for extracting information from a formatted document, comprising:***

***an input unit for inputting a formatted document (0037-0039; Maeda discloses a document obtaining unit for obtaining a structured document.).***

*a unit for analyzing the input formatted document and saving analysis results containing particular typographic information (0137-0138; 0344; 0352; 0361; Maeda discloses the important item extraction unit searches any important item contained in the received structured document and automatically extracts the important items and store them in the important item storage unit. The Examiner concludes the "important items" are the tags, which represent the structure of the document (0039-0040).).*

*a unit for identifying special character strings on the basis of the analysis results via preset values of the typographic information (0037-0040; 0214; Maeda discloses an extracting and storing unit for extracting at least one important portion from the structured document according to prescribed extraction criteria specified in advance. Using the broadest reasonable interpretation, the Examiner concludes that "prescribed extraction criteria" is equivalent to the preset values of the current invention.).*

*a unit for extracting the identified special character strings (0037-0040; 0117; 0262; 0352; Maeda discloses an extracting and storing unit for extracting and store any important items contained in the received structure document.).*

*an output unit for outputting the extracted character strings (0094; Maeda discloses a reception device for displaying the portion extracted.).*

**In regards to dependent claim 2,** *Maeda discloses the apparatus for extracting information from a formatted document according to claim 1, wherein said unit for identifying special character strings determines a certain character string as a special one on the basis of the typographic information of said formatted document when the*

*typographic information of said character string is determined to be special typographic information* (0037-0040; 0144-0146; 0153-0155; Maeda discloses a synonym database of extraction name tags, which is used is used to search for the extracted tag name.).

**In regards to independent claim 7**, Maeda discloses a *method for extracting information from a formatted document comprising:*

*inputting a formatted document, analyzing the input formatted document and saving analysis results containing particular typographic information* (0037-0039; 0137-0138; 0344; 0352; 0361; Maeda discloses a document obtaining unit for obtaining a structured document. Maeda also discloses the important item extraction unit searches any important item contained in the received structured document and automatically extracts the important items and store them in the important item storage unit. The Examiner concludes the "important items" are the tags, which represent the structure of the document (0039-0040).).

*identifying special character strings on the basis of the analysis results via preset values of the typographic information* (0037-0040; 0214; Maeda discloses an extracting and storing unit for extracting at least one important portion from the structured document according to prescribed extraction criteria specified in advance. Using the broadest reasonable interpretation, the Examiner concludes that "prescribed extraction criteria" is equivalent to the preset values of the current invention.).

*extracting the identified special character strings* (0037-0040; 0117; 0262; 0352; Maeda discloses an extracting and storing unit for extracting and store any important items contained in the received structure document.).

*outputting the extracted character strings* (0094; Maeda discloses a reception device for displaying the portion extracted.).

**In regards to dependent claim 8,** Maeda discloses *the method according to claim 7, wherein in said identifying of special character strings, a certain character string is determined as a special one on the basis of the typographic information of said formatted document when the typographic information of said character string is determined to be special typographic information* (0037-0040; 0144-0146; 0153-0155; Maeda discloses a synonym database of extraction name tags, which is used is used to search for the extracted tag name.).

**In regards to dependent claim 13,** Maeda discloses *the apparatus for extracting information from a formatted document according to claim 1, wherein the unit for identifying special character strings on the basis of the analysis results sends the typographic information to the unit for extracting the identified special character strings if the typographic information of said character strings is beyond a range of the preset values* (0037-0040; 0214; 0092; Maeda discloses an extracting and storing unit for extracting at least one important portion from the structured document according to prescribed extraction criteria specified in advance. Using the broadest reasonable

interpretation, the Examiner concludes that "prescribed extraction criteria" is equivalent to the preset values of the current invention. Maeda further discloses determining a range of the extracted important items. Thus Maeda teaches extracting the identified special character strings if the typographic information of said character strings is beyond a range of the preset values.).

**In regards to dependent claim 14,** Terayama discloses *the method according to claim 7, wherein said extracting extracts the special character strings if the typographic information of said character strings is beyond a range of the preset values* (0037-0040; 0214; 0092; Maeda discloses an extracting and storing unit for extracting at least one important portion from the structured document according to prescribed extraction criteria specified in advance. Using the broadest reasonable interpretation, the Examiner concludes that "prescribed extraction criteria" is equivalent to the preset values of the current invention. Maeda further discloses determining a range of the extracted important items. Thus Maeda teaches extracting the identified special character strings if the typographic information of said character strings is beyond a range of the preset values.).

**Note**

6. It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon

for all that it would have reasonably suggested to one having ordinary skill in the art.  
See MPEP 2123.

***Claim Rejections - 35 USC § 103***

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. **Claims 3-6 and 9-12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Maeda in view of Okamoto in view of (Pub. No.: US 2002/0065814 A1; Effective Filing Date: Jun. 30, 1998) (hereinafter "Okamota").**

**In regards to dependent claim 3, Maeda discloses *the apparatus for extracting information from a formatted document according to claim 1, wherein:***

***said formatted document is an HTML document*** (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said unit for identifying special character strings identifies a certain character string character as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said*



*character string is determined to be the biggest one among the surrounding character strings.*

However, Okamoto teaches *said unit for identifying special character strings identifies a certain character string character as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings* (0037-0040; Okamoto teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*font size*) being searched as taught by Okamoto, could have been *determined to be the biggest one among the surrounding character strings*. Thus Okamoto teaches identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamoto for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 4**, Maeda discloses *the apparatus for extracting information from a formatted document according to claim 1, wherein*

*said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.*

However, Okamoto teaches *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings* (0037-0040; 272; Okamoto teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*color and the font*) being searched as taught by Okamoto, could have been considered *to be a special one among the surrounding character strings*. Thus Okamoto teaches identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML

document when the color and the font of said character string is determined to be a special one among the surrounding character strings.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 5,** Maeda discloses *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

*said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and the font of said character string to be boldface.*

However, Okamota teaches *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the font of said character string is*

*determined to be different from the surrounding character strings and the font of said character string to be boldface* (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*boldface*) being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings*.)

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 6**, Maeda discloses *the apparatus for extracting information from a formatted document according to claim 1, wherein:*

*said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and the font of said character string to be boldface*.

However, Okamota teaches *said unit for identifying special character strings determines a certain character string as a special one on the basis of the analysis results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and the font of said character string to be boldface* (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*boldface*) being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings*.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 9,** Maeda discloses *the method according to claim 7, wherein said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis*

*results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings.*

However, Okamota teaches *said unit for identifying special character strings identifies a certain character string character as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings* (0037-0040; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*font size*) being searched as taught by Okamota, could have been *determined to be the biggest one among the surrounding character strings*. Thus Okamota teaches identifying special character strings a certain character string as a special one on the basis of the analyzing results with respect to said HTML document when the font size of said character string is determined to be the biggest one among the surrounding character strings).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 10**, Maeda discloses *the method according to claim 7, wherein said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *in said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings.*

However, Okamoto teaches *the method according to claim 7, wherein said formatted document is an HTML document, and in t said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the color and the font of said character string is determined to be a special one among the surrounding character strings* (0037-0040; 272; Okamoto teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*color and the font*) being searched as taught by Okamoto, could have been considered *to be a special one among the surrounding character strings*. Thus Okamoto teaches identifying of special character string determines a certain character string as a special one on the basis of the analysis results with respect to said HTML

document when the color and the font of said character string is determined to be a special one among the surrounding character strings.).

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 11**, Maeda discloses *the method according to claim 7, wherein said formatted document is HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *in the said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and the font of said character string is boldface.*

However, Okamota teaches *in the said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the font of said character string is determined to be different from the surrounding character strings and the font of said*



Art Unit: 2176

*character string is boldface.* (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*boldface*) being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*)

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

**In regards to dependent claim 12,** Maeda discloses *the method according to claim 7, wherein said formatted document is an HTML document* (0082; Maeda discloses the structure document can be an HTML document.).

Maeda does not expressly disclose *said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and the font of said character string is boldface.*

However, Okamota teaches *said identifying of special character string, a certain character string is determined as a special one on the basis of the analysis results with respect to said HTML document when the color of said character string is determined to be different from the surrounding character strings and the font of said character string is boldface* (0037-0040; 272; Okamota teaches means for searching the document search indexes stored in the file unit according to the query input (*character strings*). It would have been obvious to one of ordinary skill in the art that the matching character string (*boldface*) being searched as taught by Okamota, could have been *determined to be different from the surrounding character strings.*)

Therefore at the time of the invention, it would have been obvious to a person of ordinary skill in the art to combine Maeda with Okamota for the benefit of producing a structured document for display with information on the position of the document meeting the character string query (0039).

#### **Note**

9. It is noted that any citations to specific, pages, columns, lines, or figures in the prior art references and any interpretation of the reference should not be considered to be limiting in any way. A reference is relevant for all it contains and may be relied upon for all that it would have reasonably suggested to one having ordinary skill in the art. See, MPEP 2123.

***Response to Arguments***

10. Applicant's arguments have been considered but are moot in view of the new ground(s) of rejection. A new ground(s) of rejection is made in view of Maeda.

***Conclusion***

11. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).


A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to James J. Debrow whose telephone number is 571-272-5768. The examiner can normally be reached on 8:00-5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

JAMES DEBROW  
EXAMINER  
ART UNIT 2176



**Doug Hutton**  
**Primary Examiner**  
**Technology Center 2100**